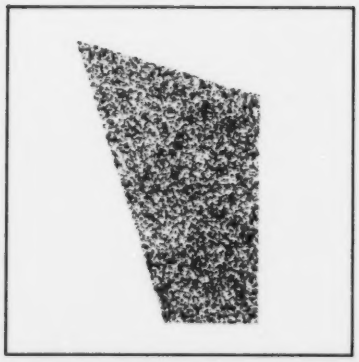
14,034

City of Toronto
Department of Public Health



Junction Triangle Historical Cohort Cancer Study

A.S. Macpherson, M.D., M.Sc.
September 1986

City of Toronto
Department of Public Health

JUNCTION TRIANGLE HISTORICAL COHORT CANCER STUDY: A STUDY OF CANCER INCIDENCE IN AN URBAN NEIGHBOURHOOD

A.S. Macpherson, M.D., M.Sc.
Medical Officer of Health

September 1986

TABLE OF CONTENTS

Acknowledgement iii

Executive Summary iv

Introduction 1

Background and Previous Studies in the Junction Triangle . 1

Design and Methods 3

Overview of the Design 3

Study Hypothesis 5

Selection of the Comparison Area 5

The Cohorts for Study 9

A. The Concept of a Cohort 9

B. Cohort Selection 9

C. Cohort Assembly 10

Linkage to the Ontario Cancer Registry 10

Certainty of Match to the Registry 11

Analysis 12

Results 14

Discussion 17

Conclusion 20

Recommendations 22

TABLE OF CONTENTS (continued)

APPENDICES

Appendix I:	TABLES 5 - 13.....	24
Appendix II:	Criteria for Certainty of Match Between Tax Assessment Identification and the Ontario Cancer Registry	33
Appendix III:	Glossary of Terms	34
REFERENCES		37

ACKNOWLEDGEMENTS

The Junction Triangle Historical Cohort Cancer Study was conducted collaboratively by the Ontario Cancer Treatment and Research Foundation and the City of Toronto Department of Public Health. Many individuals from both agencies made the study possible.

EXECUTIVE SUMMARY

The Junction Triangle Historical Cohort Cancer Study was conducted to assess the risk of cancer among residents living in an urban Toronto neighbourhood exposed to industry.

Cancer had been identified as a concern of Junction Triangle residents. However, cancer is a long-term health effect and is difficult to ascertain and study accurately in residential populations. Previous studies of the Junction Triangle had been unable to assess the possibility of long-term health effects of living in this industrialized environment. In the final report of his cross-sectional, comparative household survey of illness events and symptoms among adults and children in the Junction Triangle, Dr. Walter Spitzer recommended that a study be undertaken to "exclude the possibility that cancer adversely affects the citizens of Junction Triangle ..." The Board of Health endorsed the recommendation for a study to address this issue, September 25, 1984. In response, the present study was conducted collaboratively by the Ontario Cancer Treatment and Research Foundation and the City of Toronto Department of Public Health.

This historical cohort study identified individuals born between 1900 and 1940 who lived in the Junction Triangle between 1952 and 1956, through tax assessment rolls in Toronto City Hall. These people represented a cohort exposed to industrial emissions in their residential environment. Residents of the same age and living during the same period in the 1950's in a demographically similar, but non-industrialized area of the City were also identified. Linkage to the Ontario Cancer Treatment and Research Foundation's Ontario Cancer Registry permitted a "follow-up" search to determine cancers diagnosed between 1964 and 1982 in these two groups. Cancer incidence rates for residents of the two areas were compared. The actual (observed) cases of cancer in each area were compared to the number of cases that might have been expected if provincial rates of cancer applied in the area; thus, the ratio of "observed-to-expected" cases was used to determine if an area had an unusual or unexpectedly high rate of cancer compared to the Province of Ontario as a whole. Then, the observed-to-expected ratios for the two areas were directly compared to each other, and relative risk estimates were calculated to determine if Junction Triangle residents experienced a significantly different cancer risk than residents of the Comparison Area.

The study searched for the occurrence of cancer among 3,055 residents of the Junction Triangle and 2,911 residents of the selected Comparison Area. The two neighbourhood cohorts were comparable with regard to age and sex, but the Junction Triangle cohort was somewhat more stable and less mobile, as evidenced by their length of residence during the period of interest (1952 through 1956).

During the "follow-up" period for cancer ascertainment from the Cancer Registry (1964 through 1982), 366 cancers were observed among the Junction Triangle cohort and 342 among residents of the Comparison Area. These cases of cancer were similar in regard to age, sex and length of residence (i.e., residential "exposure") in the two cohorts. Only the certainty of matching between tax assessment information and the Cancer Registry differed between the two cohorts. A higher proportion of cancer case matches for Junction Triangle residents were classified as "definite" (53% vs. 39% for Comparison Area residents), while a higher proportion of matches for residents of the Comparison Area were less certain and classified as "possible" (42% vs. 28%).

The distribution of cancers by site (type) was nearly identical between the Junction Triangle and Comparison Area cohorts, with approximately 17% of cancers in each group being lung cancer, 12% breast cancer, 12% colorectal cancers, and 59% cancers of all other sites.

Five separate analyses were conducted:

1. Definite + Probable + Possible matches to the Cancer Registry included; all individuals not found in the Cancer Registry or all-cause mortality file were assumed alive on 31 December 1982.

2. Definite + Probable + Possible matches to the Cancer Registry included; survival of all individuals not found in the Cancer Registry or all-cause mortality file was adjusted by life expectancy for years of life remaining, according to Canadian Life Tables.

3. Definite + Probable matches to the Cancer Registry included; follow-up to 31 December 1982 (as in #1).

4. Definite + Probable matches to the Cancer Registry included; follow-up adjusted for individual life expectancies (as in #2).

5. Definite + Probable + Possible matches to the Cancer Registry included within the age-restricted cohort of residents born 1910-1934; follow-up to 31 December 1982.

   The purpose of this latter analysis was to exclude oldest and youngest cohort members. These constituted the members most likely to have died prior to the start of the follow-up period in 1964, and to have had the highest cancer incidence rates (in the case of oldest members), or were those most likely to have changed names through marriage or moved away and be unavailable to Cancer Registry follow-up (in the case of youngest members).

When all definite, probable and possible matches to the Cancer Registry were included, the cancer experience of both cohorts was generally below expected levels (below 1.00) based on provincial rates. While there appeared to be excess occurrences of cancer of the lung in Junction Triangle males (O/E ratio = 1.52) compared to provincial rates, the two cohorts did not differ significantly from each other. No differences in cancer incidence between the two cohorts were detected when the analysis adjusted for life expectancy of individual cohort members.

When the least certain matches ("possibles") were excluded from the analysis, a solitary finding of excess risk of colorectal cancers among Junction Triangle men occurred. Difficulties in matching to the Cancer Registry, greater migration and the higher rate of "possible" matches among the Comparison Area cohort resulted in under-ascertainment of cancers among Comparison Area residents. Thus, this finding is likely the result of under-ascertaining cancers among men in the Comparison Area, a problem of the study's methods. Colorectal cancers have not been associated with environmental exposures.

The historical cohort design has limitations: (1) inability to determine lifestyle, occupational or other factors which may be related to cancer risk in the historical cohorts; (2) absence of information as to the precise nature and extent of exposures to industrial emissions among individuals and the community-at-large; (3) choice of a limited period of time (1952-1956) to indicate residence and probable exposure, which serves as only a minimal estimate of people's exposures.

Comparisons of the occurrence of cancer between the Junction Triangle and Comparison Area cohorts indicated no consistent, elevated cancer risks in one cohort or the other. There was no indication that cancer incidence overall, or higher risks of cancers of the lung, colon or rectum were related to longer length of residence in either community, as measured during the period of interest for this study. The findings of this study provide no evidence of increased cancer risk related to environmental exposure in the Junction Triangle.

Environmental exposures to industrial emissions today are less than they were during the 1950's, when this study's cohort lived in the Junction Triangle. The absence of an excess cancer risk among people living in the community when industrial emissions were relatively uncontrolled, suggests that current risk to the community would be undetectable or non-existent.

Recommendations:

It is recommended:

1. That the Board of Health receive this report of the Junction Triangle Historical Cohort Cancer Study for information.

2. That no further studies to determine the risk of cancer in the Junction Triangle be planned at this time.

INTRODUCTION

The last ten to fifteen years have seen increasing concern raised by the lay and medical communities over the potential adverse effects of industrial pollution in residential neighbourhoods. Aside from the effects on property values and vehicular traffic, the conjunction of industrial and residential urban areas has aroused concern over the health of residents exposed to emissions from industrial plants. Residents want more information about both the short- and the long-term effects on themselves and on their children.

The study described here was carried out to address one specific long-term health issue by assessing the risk of cancer in such an exposed population - the Junction Triangle neighbourhood in Toronto.

This study was done as a collaboration between the Ontario Cancer Treatment and Research Foundation and the City of Toronto Department of Public Health.

BACKGROUND AND PREVIOUS STUDIES IN THE JUNCTION TRIANGLE

The Junction Triangle is located in the northwest area of the City of Toronto. It corresponds to one complete census tract (#98); the land surface area is 0.70 square kilometres and it is bounded by railway lines to the east, north and west.

The 1976 Land Use data from the Metro Toronto Planning Department showed that 19 percent of this area was zoned for industrial use (1). There are 20 industries in the area including seven major manufacturing plants which produce mainly paint, resin and glue products. An important group of chemicals widely used in area industries is organic solvents.

The history of the Junction Triangle reaches back over one hundred years when the area was the junction of two major roads. Because of the local clay deposits, it became an area for houses and brick-making rather than farming. When the railway lines were laid and a railyard established, homes and jobs were created and a real estate boom followed. As the transport routes created an ideal location for industry, and the opportunity of walking to work was considered an advantage (2), the growth of industry and residences side by side set the pattern for development in the Junction Triangle area. In contrast to the past, today only nine percent of Junction Triangle residents work within the census tract.

In the 1970's, citizen groups organized to express their concerns about the environment. Two groups of concerned Junction Triangle residents were formed to investigate noxious odours from local industry. The Department of Public Health became involved in the Junction Triangle in the early 1970's, in response to citizens' concerns (3).

In the summer of 1979, a resident group (the Junction Triangle Anti-Pollution Committee) conducted a non-random survey of Junction Triangle residents. Among 547 households, at least 30 percent of residents complained of frequent colds, unusual coughing, sneezing, eye irritation, skin rashes, breathing difficulties or unusual tiredness. Two percent suffered from cancer. The study reported that "49 percent of respondents feel that their illnesses are related to pollution" (4).

A major chemical spill of vinyl acetate and butyl acrylate in April 1982 prompted Toronto City Council to establish a Mayor's Task Force (5) and to engage an outside consultant to investigate the residents' health concerns.

The Mayor's Task Force established the Environmental Liaison Committee with membership from the community, industry and government. This committee has met bi-monthly since its inception and acted as a forum for discussion among residents, the industries in the Triangle and government officials.

Dr. W.O. Spitzer from McGill University was commissioned by the Task Force in 1983 to conduct an epidemiological study of Junction Triangle residents. The report (the Spitzer Study) was submitted to the Board of Health in April 1984, and was released in final form in September 1984. The results showed that children in the Junction Triangle reported more eye, nose, throat and skin symptoms in the previous two weeks and in the previous year than children from a comparison area of the City (6). Dr. Spitzer recommended that a comprehensive health promotion program be undertaken. The Department of Public Health conducted the Junction Triangle Health Assessment Program in the spring of 1985, in which all children 14 years of age or younger who lived or attended school in the Junction Triangle were invited to participate. The results of this program were reported to the Board of Health in November 1985. The conclusion was that, of the health concerns assessed (eye, nose and throat irritation, skin irritation and respiratory function), there was no evidence of increased health problems in Junction Triangle childrer (7).

Dr. Spitzer stated in his report: "There were very few cases of cancer, but our sampling techniques were inadequate to study cancer and no conclusions can be reached in this regard" (6).

He recommended that:

> "The governments of the Province of Ontario and of the City of Toronto should commission the Ontario Cancer Treatment and Research Foundation to undertake a study to exclude the possibility that cancer adversely affects the citizens of Junction Triangle and neighbouring areas compared to Toronto as a whole."

At its meeting on September 25, 1984, the Board of Health endorsed this recommendation. The design and methodology of the historical cohort study was received by the Board of Health at its special meeting on December 5, 1984 (8).

DESIGN AND METHODS

OVERVIEW OF THE DESIGN

Cancer had been identified as a concern of Junction Triangle residents, but previous studies of the area have been unable to assess the possibility of long-term health effects of living in this industrialized environment. Therefore, an historical cohort study was designed to determine the occurrence of cancer in residents of the Junction Triangle. As part of the study, cancer incidence among people known to have lived in the Junction Triangle between 1952 and 1956 was compared to the incidence of cancer among residents of a similar, but non-industrialized area of the city.

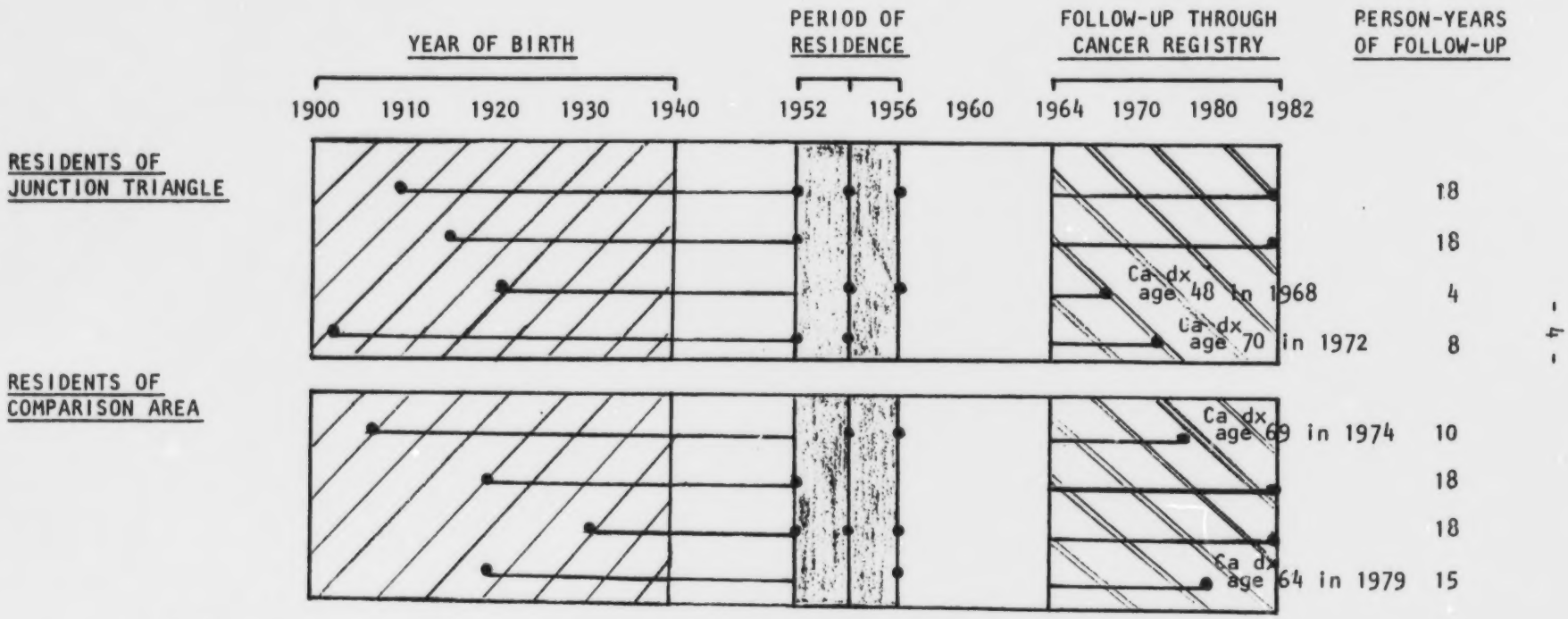
Individuals born between 1900 and 1940 who lived in the Junction Triangle between 1952 and 1956 were identified through the tax assessment rolls in Toronto City Hall. These people represented a cohort exposed to industrial emissions in their residential environment. Residents of the same age and living during the same period in the 1950's in a demographically similar area of the city with no industry were also identified. Linkage to the Ontario Cancer Treatment and Research Foundation's Ontario Cancer Registry permitted a "follow-up" search to determine cancers diagnosed between 1964 and 1982 in these two groups.

Cancer incidence rates for residents of the two areas were examined in several ways. The actual (observed) cases of cancer in each area were compared to the numbers of cases that might have been expected if provincial rates of cancer applied in the area; thus, the ratio of "observed-to-expected" cases was used to determine if an area had an unusual or unexpectedly high rate of cancer. Additionally, the "observed-to-expected" ratios for the two areas were compared to each other to determine if Junction Triangle residents experienced a significantly different cancer risk than residents of the comparison area. Then, relative risk estimates, with 95% confidence intervals, were calculated. Relative risk, which is the ratio of cancer incidence rates in the two areas, determines the risk of cancer among Junction Triangle residents compared to residents of the industrially unexposed area. An overview of the study design is shown in Figure 1, which illustrates the residency and follow-up experience of four individuals in each cohort, as examples. In addition, key epidemiologic concepts used in this report are explained in the Glossary of Terms (Appendix III).

FIGURE 1

HISTORICAL COHORT STUDY DESIGN

(WITH EXAMPLES OF THE RESIDENCY AND FOLLOW-UP EXPERIENCE OF FOUR INDIVIDUALS IN EACH COHORT SHOWN)



LEGEND

Ca dx = Diagnosis of Cancer

= Years of birth, 1900-1940

= Years of residence in area, 1952-1956, determined through Tax Assessment files

= Years of follow-up through Cancer Registry, 1964-1982

STUDY HYPOTHESIS

The null hypothesis to be tested in this study was that the risk of developing cancer was the same for residents who lived in the Junction Triangle between 1952 and 1956, a period of relatively uncontrolled industrial emissions, and for those who lived at the same time in a demographically similar, but non-industrialized area of Toronto.

Put another way, the study seeks to answer the question as to whether residents who lived in the Junction Triangle between 1952 and 1956 were at greater risk of developing cancer than those who lived in a demographically similar, non-industrial area of the City.

SELECTION OF THE COMPARISON AREA

A cohort for comparison with Junction Triangle residents comprised residents of another Toronto census tract similar in demographic characteristics, but with little evidence of industrial activity in the 1950's. The analysis which enabled selection of a comparison area was conducted by Dr. Robert Murdie, Associate Professor of Geography, York University (9).

Industrial characteristics of the City were determined from historical land use data. Land use data for minor planning districts (i.e., aggregations of census tracts) for 1958 were examined, together with similar data for 1976, by which time information was provided on a census tract basis (1,10). The proportion of an area's land mass devoted to industry was calculated. Areas relatively free from industry were defined as those in which less than 10% of the total land area was occupied by industry.

In general, industrial activity in the City of Toronto has declined since the early 1950's. In addition, there was considerable variation in the degree of industrialization within each minor planning district. Because land use data were unavailable on a census tract basis in the 1950's, it was assumed that each area had sustained more industrial use of its land in the early 1950's than in the 1970's and 1980's. The minor planning district encompassing the Junction Triangle declined from 22% of its total area in industrial use in 1958 to 12.3% by 1976; industrial land use in the Junction Triangle census tract, specifically, was 19.1% in 1976.

Factor analysis was used to analyze 48 sociodemographic variables on the 1951 census population in each of Metropolitan Toronto's 220 census tracts, yielding a measure of demographic match between the Junction Triangle and every other census tract in Metro. This statistical method reduces information into related clusters. The variables chosen for inclusion in the analysis were measures of

such population characteristics as age and family structure, household size, ethnicity, income, occupation and educational attainment. Five major sociodemographic factors resulted from the analysis, accounting for 75% of the variance represented by the original 48 variables. The factors summarized a census tract's population along the following lines:

- Socioeconomic status: managerial/professional vs. blue collar occupation
- Socioeconomic status: clerical vs. service or unemployed occupation
- Ethnicity: Anglo-Saxon vs. non-Anglo-Saxon
- Family status: Family vs. non-family household status
- Residential turnover: high mobility vs. low mobility households

Standardized scores on each of the five factors were computed for each census tract and demographic similarity between tracts was measured by squared distances in factor scores.

Census tracts most demographically similar to the Junction Triangle tended to be those in closest proximity to it. It was judged desirable to choose a comparison area at some distance from the Junction Triangle so that residents were unlikely to have been exposed to airborne emissions from Junction industries. Another ten census tracts judged similar in demographic characteristics were within minor planning districts having more than 10% of land area in industrial use in 1958; these were excluded as unsuitable comparison areas.

The selected comparison census tract was suitably non-industrialized, with 5.8% industrial land use in 1958 and no industry by 1976. However, the area differed somewhat from the Junction Triangle in ethnic composition (see Table 1); a higher proportion of its population was of Anglo-Saxon origin - 69.7% compared to 55.4% of the Junction Triangle's population. In 1951, a sizable proportion of Junction Triangle's population was of Eastern European ethnicity: 21% Ukrainian and 6% Polish.

Both the 1951 Census and the 1952-1956 tax assessment rolls indicated that the population of the census tract selected as the Comparison Area was substantially smaller than that of the Junction Triangle. Therefore, it was decided to expand the Comparison Area, by including streets which were contiguous to the original census tract but beyond its boundaries. Because expansion of the area extended beyond original census tract boundaries into two adjacent tracts, precise information on the demographic characteristics of the population in this expanded area was not available.

However, factor analysis scores on the sociodemographic characteristics of the two census tracts from which the additional streets were drawn, suggested these tracts were similar to the selected Comparison Area tract. They, too, differed from the Junction Triangle in having a higher proportion of their populations of Anglo-Saxon origin.

This small expansion of the geographic boundaries of the Comparison Area yielded a population of nearly equivalent size in the two areas. Given the provincial cancer incidence rates and the person-years of follow-up in the study, the populations were large enough to detect a two-fold increase in cancer risk between the cohorts (i.e., a relative risk of 2.0), with an 80% power to detect such a level of increased risk and a five percent probability that such a difference, if found, is due only to chance (11).

TABLE 1

DEMOGRAPHIC CHARACTERISTICS OF JUNCTION TRIANGLE AND COMPARISON AREA, 1951

Area/Population Characteristic	Junction Triangle	Original Comparison Census Tract *
Land Area	0.70 sq.km	0.67 sq.km
Total number residents	5,392	1,850
Age Distribution		
0 - 19	23.3%	20.1%
20 - 64	64.2%	71.0%
65+	12.5%	8.9%
Median Income		
Males	$2,300	$2,200
Females	1,500	1,300
Official Language		
English Only	94.6%	92.2%
Other	5.4%	7.8%
Country of Origin		
British Isles	55.4%	69.7%
Italian	5.5%	2.0%
Other European	30.8%	16.7%
Other	8.3%	11.6%

Source: 1951 Census, Statistics Canada

* The Comparison Area was expanded beyond the boundaries of the original census tract to include a few contiguous streets from two adjacent census tracts. Demographic information is not available on partial census tracts. Demographic characteristics of the two complete, adjacent tracts indicated the areas were similar to the original census tract chosen. The original census tract probably well represents the demographic character of the enlarged Comparison Area.

THE COHORTS FOR STUDY

A. THE CONCEPT OF A COHORT

The interest of this study is examining cancers among a group of people (a cohort) with a common exposure, that is, people who lived in the Junction Triangle during a period of relatively uncontrolled industrial emission. Within this cohort, individuals undoubtedly will have had varying lengths and degrees of exposures. The "exposed" cohort is compared to an "unexposed" cohort with different residential exposure, a group of individuals who lived at the same time in a city neighborhood with little industry.

B. COHORT SELECTION

In order to relate the development of cancers to exposures in the Junction Triangle, study cohorts were selected based on residence in the Junction Triangle and the Comparison Area over a particular period of time and for varying numbers of years.

The cohorts for each area were made up of all residents listed on the tax assessment files who were born between the years 1900 and 1940, and who lived in the respective areas between 1952 and 1956. Because children are not listed by name on the tax assessment files, they were not included for study. Residency was determined by actual occupancy of a residential dwelling at a particular address, and included resident-owners or rentors of an address. Therefore, these criteria for cohort inclusion ensured as complete an enumeration as possible of the adult residential population of the two areas.

Excluded from the cohorts were those with unknown birth year or sex. Also excluded were individuals known to have died prior to the beginning of the follow-up period (i.e., prior to 1 January 1964). A few individuals were discovered to have lived in both study areas; they were excluded from study as well. These procedures resulted in 3,055 residents of the Junction Triangle and 2,911 Comparison Area residents comprising the cohorts for study.

Residents of interest were those presumed to be exposed to industrial pollutants by virtue of residency in the Junction Triangle at any time during the five-year period between 1952 and 1956. The cohort of comparison was individuals residing in the selected Comparison Area of the city during the same time period. This historical time period was chosen for several reasons. The mid-1950's was a time of relatively uncontrolled industrial emissions in the Junction Triangle area. People born in the study cohort period (1900-1940) and living during the mid-1950's would be in the age groups of highest cancer risk by the follow-up period (1964-1982). In addition, the follow-up period was

sufficiently far in the future to permit the detection of cancers in the selected population. Thus, if people were exposed to potential risk factors for cancer between 1952 and 1956 any related cancers that subsequently developed would be most likely to be detected between the mid-1960's and 1980's.

C. COHORT ASSEMBLY

A City of Toronto computerized streets inventory, listing all streets by census tracts, was used to identify all addresses within each of the two study areas. The city's tax assessment files were searched for specific streets to determine residents with birthyears between 1900 and 1940, living in each area during the period of interest. The tax assessment files contained information about residents for taxation purposes one year prior to the actual taxation year. The years of information collection (1952, 1954, 1956) were used to establish residency at a particular address.

Information gathered from the tax assessment files for each individual resident included: name, address, occupation, birthdate/alternate birthdate if this varied from year to year on the assessment files, spouse's birthdate, census tract and sex. Sequential searching of the tax assessment files for the years 1952, 1954 and 1956 indicated the individual's presence in the area in each of these years, and served as an indicator of length of residence during the period of interest.

Length of residence was taken as an approximate indicator of "exposure" for the purposes of this study. Those people whose names appeared on the assessment rolls in only one of the three years were considered as having one year exposure during the period of interest; those whose names appeared in two consecutive rolls, either 1952 and 1954, or 1954 and 1956 were considered to have two years exposure; residents whose names appeared on all three assessment rolls examined were assigned three years' exposure. Finally, the 15 people whose names appeared on the 1952 and 1956 rolls but not on the 1954 rolls, were assigned to the lowest, (one year) exposure level.

Thus, a surrogate measure of non-employment industrial exposure was generated. The estimates of exposure were minimal estimates only, and served as a crude guide to possible exposure. While lifetime exposure data, both at home and in the workplace, would be preferable, it was not feasible to obtain such detailed data, given the historical nature of this study.

LINKAGE TO THE ONTARIO CANCER REGISTRY

Names of cohort members and demographic information were collected in batches of 500 and relayed to the Ontario Cancer Treatment and

Research Foundation for linkage with the Ontario Cancer Registry. Each batch of 500 consisted of approximately equal numbers of Junction Triangle and Comparison Area residents.

Names of cohort members were searched in the cancer incidence files of the Ontario Cancer Registry. This linkage between the cohorts and the Registry was carried out by health records technicians familiar with the process of name-searching and with the records of the Registry. Due to extensive media coverage of the issues, the technicians were aware of the concern in the exposed neighbourhood; however, they were unaware of the cohort membership of a given individual. Thus, the searching for cancer diagnoses was carried out in a "blind" fashion.

At the time that the linkage was done, the Ontario Cancer Registry included data on all new cancers diagnosed, by hospital or pathology report, in the province of Ontario between the years 1964 and 1982. In addition to searching these records, the health records technicians searched the file of Ontario deaths from all causes (1964-1982) and the file of reports from the cancer treatment centres in the province. (These latter two sources of data were not integrated into the database of the Ontario Cancer Registry at the time the searching was carried out.)

CERTAINTY OF MATCH TO THE REGISTRY

Linkage between the cohorts and the Ontario Cancer Registry resulted in the identification of matches with varying degrees of certainty. The Registry contains much identifying data; the historical tax assessment rolls, on the other hand, contained only limited information about the individuals. This information was confined largely to name, age, and address of residence. Therefore, it was not possible to match all individuals with the same degree of certainty that the match was a correct linkage between records on the same person. The possibility existed that the match was a linkage of records on two different people with the same name and the same age, a situation that was quite likely if the name was a common one. As a result, a linkage hierarchy was developed to reflect the degree of certainty in the matching. A match between a cohort member and the Registry was defined as definite, probable, possible, or non-match.

A match was definite if the surname was uncommon and the date of birth agreed within two years, or if the Registry record showed that the patient was treated or died in the metropolitan region. Probable matches arose for a number of reasons. A match was probable if the surname was uncommon and the Registry had only the year of birth, or the Registry had only a partial given name. In addition, a match was probable if the surname was uncommon but there was partial agreement on the birthdate, or if the patient was treated or died outside of the metropolitan region but the birthdate or given names were in agreement. A match was also

probable if the surname was common, but additional information was in agreement, such as complete birthdate, given names, or maiden name. Finally, a match was possible if the surname was common and any of the following applied: the patient was treated or died outside of Ontario, the middle name or initial disagreed, the Registry had only year of birth or partial given name, or the birthdates did not agree within two years. Also, a match was possible if two or more persons of the same name were found in the Registry with a date of birth within two years of that recorded on the assessment roll. An attempt was made to convert the possible and probable matches to definite matches, by searching both older and more recent assessment rolls for further information, such as complete date of birth and occupation, and by searching death certificates for address and occupation. In the event that no clarifying information could be found, possible and probable matches remained, and were coded as such for the analysis. Criteria for certainty of match are summarized in Appendix II.

ANALYSIS

Information obtained from the Ontario Registry on a match included the site of cancer, date of diagnosis, vital status, and cause of death, if applicable. Those individuals not found in the Registry, or in the all-causes mortality file, were assumed to be alive and cancer-free at the end of the follow-up period, 31 December 1982. Cohort analysis was carried out, with the person-years approach, using the computer programme MANYEARS developed by J. Peto and modified by M. Coleman et al (12). Annual sex- and age-specific Ontario cancer incidence rates were used to generate the expected numbers of cancer diagnoses in the two cohorts. Differences between the observed-to-expected ratios of the two cohorts were tested for statistical significance using an asymptotic chi-squared test for homogeneity (13). Statistical significance was set at an alpha level of 0.05. Relative risk estimates were calculated as a means of directly comparing the cancer experience of the two cohorts. The cancer incidence rate of each cohort (i.e., the number of cases of cancer divided by person-years contributed by the group under analysis) was compared as a ratio, and 95% confidence intervals around the risk estimate determined (14).

The comparisons of interest were those between the two cohorts, rather than between each cohort and the provincial cancer incidence rate. This latter comparison could be affected by many factors, including urban/rural differences in cancer rates. The cohorts, however, were matched as nearly as possible on various characteristics which could affect the residents' risk of cancer of different types.

Five separate analyses were carried out, and are summarized in Table 2. In the first analysis, definite, probable and possible matches were included in the observed figures, and individuals for

whom no follow-up (i.e., cancer match) was obtained were assumed to be alive on 31 December 1982. In the second analysis, again, definite, probable and possible matches were included in the observed figures, but the follow-up of individuals not in the Registry or mortality files was adjusted by the application of the Canadian Life Tables for 1955-1957 (15). The life table provided the expected number of years of life remaining; for a given cohort member, person-years of follow-up were corrected by that life expectancy. For example, a 54 year-old man in 1955 had a life expectancy of 21 years. Person-years of follow-up, then, would be only 12 years (i.e., 1976 minus 1964), rather than the 18 years (i.e., 1982 minus 1964) calculated in analysis #1. This method of adjusting the follow-up was deemed desirable, since those who died or moved out of the Province prior to 1964 would erroneously inflate the number of person-years. In the subsequent analyses (#3 and #4), only definite and probable matches were included, with the third analysis being the conventional follow-up to the last date possible (i.e., 31 December 1982), and the fourth analysis being adjusted for the individual life expectancies. In these two analyses, individuals who were classified as only "possible" matches to the Registry were assumed alive to the respective follow-up dates and, thus, contributed person-years to the denominator.

TABLE 2

ANALYSIS OF DATA BY CERTAINTY OF MATCH AND FOLLOW-UP METHOD

ANALYSIS NUMBER	CERTAINTY OF CANCER MATCH	FOLLOW-UP DATE
1	Definite + Probable + Possible	31 December 1982
2	Definite + Probable + Possible	Life-Table Adjusted
3	Definite + Probable	31 December 1982
4	Definite + Probable	Life-Table Adjusted
5	Definite + Probable + Possible among age-restricted cohort of reidents born 1910-1934	31 December 1982

It should be noted that the effect of analyses #2 and #4 was to diminish the number of follow-up years, while the effect of analyses #1 and #2 was to increase the number of observed cancers, through the inclusion of possible matches. Analysis #2 would yield the largest observed-to-expected ratios.

The fifth analysis involved examination of the Junction Triangle and the Comparison Area residents, restricting the cohorts to those born between 1910 and 1934 (i.e., excluding the oldest and the youngest segments of each cohort). The purpose of eliminating the oldest and youngest cohort members was to exclude those individuals for whom follow-up might have been particularly incomplete. The oldest cohort members were those who:

1. were most likely to have died prior to the beginning of the follow-up period in 1964;

2. contributed the most person-years to those analyses which assumed individuals to be alive and disease-free through 1982; and

3. would have the highest cancer incidence rates.

Any error in ascertaining the last cancer-free date of this subgroup would be magnified by these factors, and could bias the overall results. With regard to the youngest cohort members, of whom there were very few, it was believed that follow-up would be difficult because of (a) high mobility in the young; and (b) name changes by many of the young women. The restricted cohorts were analyzed in the same way as were the complete cohorts.

The cancer sites chosen for detailed analysis, in addition to all cancer sites combined, were: cancers of the trachea, bronchus, and lung; female breast cancer; cancers of the colon and rectum. These sites were selected for several reasons. Cancers of these sites occurred frequently enough in the two cohorts to allow for examination by sex and length of residence. They are cancers with major public health implications. Together, these cancer sites cover a range of organ systems and have different etiologies, as far as is known. While it might have been interesting to examine other, rarer types of cancer in the two cohorts, it was not possible to do so because of their rarity. Then too, any deviations from expectation in the occurrence of these cancers could be assessed by looking at the data for all cancer sites combined, where an excess in the incidence of the rarer cancers might be evident.

RESULTS

A total of 5,966 subjects were available for analysis, 3,055 (51%) residents of the Junction Triangle (JT), and 2,911 (49%) residents of the Comparison Area (CA), contributing a total of 52,148 (JT) person-years and 49,911 person-years (CA) respectively. Table 3 displays the basic demographic and matching information on the two cohorts. The cohorts were comparable in their distributions of sex and age. With regard to minimal length of residence in the respective area, the indicator chosen to approximate exposure, the Junction Triangle cohort was somewhat more stable than the

Comparison Area. Just over 28% of the Junction Triangle residents were identified as living in the census tract in all three assessment years, while only 19% of the Comparison Area residents were so identified.

TABLE 3

DEMOGRAPHIC, EXPOSURE AND FOLLOW-UP CHARACTERISTICS BY STUDY COHORT

		JUNCTION TRIANGLE		COMPARISON AREA	
		N	(%)	N	(%)
TOTAL COHORT SIZE		3055	(100.0)	2911	(100.0)
SEX					
Male		1,555	(50.9)	1,417	(48.7)
Female		1,500	(49.1)	1,494	(51.3)
YEAR OF BIRTH	(AGE IN 1982)				
1900-1909	(73-82)	804	(26.3)	751	(25.8)
1910-1919	(63-72)	918	(30.1)	860	(29.5)
1920-1934	(48-62)	1,314	(43.0)	1,272	(43.7)
1935-1944	(38-47)	19	(0.6)	28	(1.0)
EXPOSURE	(MINIMUM LENGTH OF RESIDENCE)				
1 year		1,584	(51.8)	1,767	(60.7)
3 years		611	(20.0)	605	(20.8)
5 years		860	(28.2)	539	(18.5)
PERSON-YEARS		52,148 years		49,911 years	

The distribution of the observed cancers by cohort membership and demographic variables is shown in Table 4. The certainty of matching with the Cancer Registry differed between the two cohorts. A greater proportion of the Junction Triangle matches than the Comparison Area matches were "definite" (53% vs. 39%), while a smaller proportion of Junction Triangle matches than Comparison Area matches were "possible" (28% vs. 42%). The proportion of matches which were "probable" was the same in the two cohorts (19%). The observed cancers formed the same age, sex, and length of residence distributions in each of the two cohorts.

TABLE 4

DISTRIBUTION OF OBSERVED CANCERS BY COHORT, CERTAINTY OF REGISTRY MATCH, SEX, YEAR OF BIRTH, AND EXPOSURE

	JUNCTION TRIANGLE		COMPARISON AREA	
	N	(%)	N	(%)
TOTAL OBSERVED CANCERS	366	(100.0)	342	(100.0)
REGISTRY MATCH				
Definite	193	(52.7)	133	(38.9)
Probable	69	(18.9)	64	(18.7)
Possible	104	(28.4)	145	(42.4)
SEX				
Male	201	(54.9)	170	(49.7)
Female	165	(45.1)	172	(50.3)
YEAR OF BIRTH				
1900-1909	131	(35.8)	110	(32.2)
1910-1919	124	(33.9)	116	(33.9)
1920-1934	111	(30.3)	114	(33.3)
1935-1944	0	(0.0)	2	(0.6)
EXPOSURE				
1 year	166	(45.4)	184	(53.8)
3 years	75	(20.5)	78	(22.8)
5 years	125	(34.1)	80	(23.4)

The cancer site distributions (APPENDIX I, Table 5) between the two cohorts were nearly identical, with approximately 17% of cancers in each cohort being lung cancer, 12% breast cancer, 12% colorectal cancers, and 59% other cancers.

Tables 6 - 9 (APPENDIX I) display the observed and expected numbers of incident cancers in the two cohorts, relative risks and 95% confidence intervals around the risk estimate, for different cancer sites and for the two sexes. If the possible registry matches are included in the analysis, along with the definite and probable matches (Tables 6 and 7), then there is little deviation of observed from expected, and the observed-to-expected ratios are

very close to 1.00. In general, the number of observed cases of cancer is lower than expected, with the notable exceptions of cancer of the lung in Junction Triangle males (O/E = 1.15), and cancer of the lung in the Comparison Area females (O/E = 1.52), where there appear to be excess occurrences of cancer compared to provincial rates. These ratios, however, are not statistically significantly different from those in the other cohort (Table 6). The lower bounds of the confidence intervals are below 1.00, thus, this direct comparison between the cohorts does not show significantly elevated risk in the Junction Triangle cohort. Adjusting the person-years for the life expectancy of individual cohort members had little effect on the overall results (Table 7).

Excluding the possible matches, however, had a pronounced effect: while all observed-to-expected ratios are substantially decreased, there is a greater decrease in the Comparison Area than in the Junction Triangle (Tables 8 and 9). In these analyses, the observed-to-expected ratios and relative risks for colorectal cancers and total cancers among males were higher for men in the Junction Triangle. The significance of the ratio for all sites was largely accounted for by colorectal cancers. Exclusion of the possible matches led to a marked under-ascertainment of incident cancers in both the Junction Triangle and the Comparison Area cohorts. However, the excess of possible matches in the Comparison Area cohort lowered the observed-to-expected ratios for this cohort more than in the Junction Triangle cohort, particularly with regard to colorectal cancers. The analyses of all cancers, lung cancers and colorectal cancers by length of residence (Tables 10 - 12), showed no trends suggesting that cancer incidence was related to length of residence in the respective community, as measured during the period of interest for this study.

Restricting the cohorts to those born between 1910 and 1934, thus eliminating oldest and youngest cohort members, (Table 13), indicates the relatively large number of cancers contributed by oldest cohort members in the full analysis. Relative risks remained essentially unchanged by this age restriction.

DISCUSSION

There are, in historical studies such as this, limitations imposed both by the study design and by the availability of records. One problem in studies of this type is the inclusion in the cohort of people who have no possibility of being diagnosed with cancer during the follow-up period. This may present itself in two ways: as people who moved away from the province or as people who have died prior to the start of follow-up (1964).

The Ontario Cancer Registry is Province-wide. It will allow tracking the cancer experience of individuals who move elsewhere in the Province, but will not allow identification of cancers in

people who have moved out of Ontario. Thus, mobility of the cohorts may affect detection of cancers. This study could not account for possible differences in migration patterns between the two cohorts. In fact, the tax assessment file searches suggested a lesser degree of mobility among Junction Triangle residents, as indicated by a higher proportion of this cohort resident in the area throughout the five-year residency "exposure" period. Members of the Comparison Area cohort may have been somewhat more likely to move; thus, cancers among them may be under-estimated. Lower migration, as evidenced among the Junction Triangle residents, generally enables better "follow-up" through the Cancer Registry and higher levels of cancer ascertainment. The effect of this study's inability to control for migration among the cohorts is likely to have been an under-estimation of cancers in the Comparison Area cohort compared to the Junction Triangle cohort.

The analysis of the restricted cohort in this study (i.e., those born 1910-1934) was an attempt to compare findings of cancer among those in the age groups least likely to have died prior to 1964. This analysis was comparable to the complete cohort analysis, as indicated by relatively unchanged relative risk estimates between the analyses, suggesting that loss of members due to early deaths undetected by this study did not differ markedly between the cohorts.

Children resident in an area are not identified on tax assessment rolls. Inclusion for follow-up and cancer ascertainment in this study was limited to adults born before 1941. Therefore, the study was unable to detect cancers in young children.

Given the paucity of identifying information on the members of the two cohorts, it was inevitable that there might be under-ascertainment of cancer occurrence in the cohorts. While the use of the "possible" registry matches is problematic from the point of view of correctness of the match, the exclusion of the possible matches led to an obvious under-ascertainment of cancer cases in both cohorts. This under-ascertainment, however, affected the Comparison Area more than it did the Junction Triangle, probably because of differences in the ethnic distributions of the two cohorts. It is impossible to quantify this ethnic difference within the two cohorts selected for study. However, it was clear, from the distribution of surnames in the two cohorts, that the presence of "unusual" names in the Junction Triangle facilitated the matching of subjects to the Cancer Registry; the presence of relatively more common surnames in the Comparison Area, due to a higher proportion of the population being Anglo-Saxon origin, rendered the matching process more difficult. The factor analysis procedure for selection of a demographically similar area for comparison attempted to account for differences in ethnicity; however, the constraint that the neighbourhood to be used for comparison had to be non-industrial, forced the selection of an imperfectly-matched comparison area. The study was also unable to control for known risk factors for some of the cancers detected,

such as smoking history, reproductive history, or occupational exposure. Differences between residents of the two communities with regard to such factors which influence the risk of cancer, was handled somewhat by comparing each cohort, separately, to the provincial cancer experience. In general, both cohorts had cancer rates at or below provincial expectation.

Data from the Toronto Community Health Survey, 1983, were examined relative to smoking and alcohol consumption habits in an attempt to explore factors that might be related to lung and colorectal cancers (16). Country of birth, asked in the Community Health Survey, serves as the only indicator of ethnicity. It was known from the 1951 census data that the Junction Triangle had a higher proportion of residents of Eastern European origin, while a higher proportion of the Comparison Area population was of Anglo-Saxon origin. Current Toronto residents of these two ethnic groups, as represented in the Community Health Survey, did not differ significantly with regard to smoking habits or amount of alcohol usually consumed. However, these data are limited in their ability to help explain this study's findings, as the Community Health Survey reflects current habits in the population; habits that might influence a health outcome such as cancer would have to take account of people's behaviour 20 to 30 years ago.

The choice of a suitable comparison for the cancer rates in the Junction Triangle cohort is a critical one. Tests of the significance of individual observed-to-expected ratios (based on the Poisson distribution (17)) showed that, in comparison with provincial cancer rates, both the Junction Triangle and the Comparison Area cohorts experienced an excess of cancer occurrence with regard to lung cancer (Junction Triangle males and Comparison Area females) and to cancers of all sites combined (males in both cohorts). In the Junction Triangle cohort, the excess of all cancer sites combined is nearly completely accounted for by the excess in lung cancer. While these deviations from the provincial expectation are interesting, the comparison between cohort and province is not an appropriate one, given the lack of adjustment for many of the factors which affect cancer incidence. There is a well-established excess of cancer associated with urbanization, particularly for such cancers as lung and rectum (18). Data from the Ontario Cancer Registry support this association (19). The urban-rural difference alone could account for the deviation of each cohort from the provincial expected frequency. Thus, the most useful comparison to be made was a direct one between the two urban cohorts.

The direct comparisons between the cohorts did not indicate consistent, elevated cancer risks with regard to sex or cancer site in one cohort or the other. The significant difference in the observed-to-expected ratio and the elevated relative risk for colorectal cancers in men occurred when "possible" registry matches were excluded from the analysis. The apparent emergence of a risk of colorectal cancers alone, among men, was likely due

to the very low observed-to-expected ratio for this cancer among men in the Comparison Area. The difficulties in matching, and the higher rate of possible (and, therefore less certain) matches among the Comparison Area cohort probably resulted in under-ascertainment of cancers among that cohort. Thus, this isolated finding is likely a result of the under-ascertainment bias present in this method. In addition, this difference is unlikely to be attributable to exposure to some carcinogen(s) among those who lived in the Junction Triangle; colorectal cancers have not been associated with environmental exposures (20).

This study's sample sizes were not large enough to detect differences between the cohorts in regard to rarer cancers with sufficient power. However, examination of the distribution of cancers by type in each cohort does not indicate an excess incidence of groups of cancers which might feasibly be related to environmental exposures in either cohort.

Analysis of the age-restricted cohorts of those born between 1910 and 1934 revealed the same lack of excess cancer incidence or excess risk in the Junction Triangle as compared with that in the Comparison Area. The analysis also indicated that there was some over-estimation of the number of person-years in both the Junction Triangle and the Comparison Area cohorts. The age-restricted cohorts had observed-to-expected ratios more nearly in line with what would have been expected by chance alone, whereas the complete cohorts had ratios which indicated under-ascertainment of cancer cases or over-ascertainment of person-years at risk. The oldest segment of each cohort would have contributed the most to any over-estimation of the expected numbers because of the very high cancer incidence rates which would be applied to the person-years at risk in these sub-groups.

Risk of cancer appeared unrelated to duration of residence during the period of interest measured in this study. The creation of an exposure "window," a selected period of time during which evidence of residency in an area was considered as a measure of potential chemical exposure, is necessarily arbitrary. The time period 1952 through 1956 was chosen as a time of relatively uncontrolled industrial emissions. Setting an arbitrary "window" meant that some people would have lived in an area before and/or after this time, and their true length of residence was not captured by this study. The estimates of exposure used here were only minimal estimates. In addition, the true chemical nature of exposures remains unknown, as historical records of chemical use do not exist.

CONCLUSION

This study has compared cancer incidence in the Junction Triangle and the Comparison Area, first to the cancer experience of the Province of Ontario, and then directly to each other as a means of

assessing the risk of cancer among residents of the Junction Triangle community who are exposed to industry in their residential environment. In general, the cancer experience of both areas was below expected levels based on provincial rates.

There is no evidence of increased cancer occurrence related to environmental exposure in the Junction Triangle compared with that in the Comparison Area, as determined in the cohort of residents living in those neighbourhoods between the years 1952 and 1956. The cancers which occurred in the two cohorts were not associated with the length of time the cohorts' members were residents of these areas. The solitary finding of excess risk of colorectal cancers among Junction Triangle men occurred when "possible" matches to the Cancer Registry were excluded from the analysis. Difficulties in matching to the Cancer Registry, greater migration, and higher rates of "possible" (less certain) matches among the Cómparison Area cohort are likely to have contributed to under-ascertainment of cancers among Comparison Area residents. Thus, this finding is likely the result of under-ascertaining cancers among Comparison Area men, a problem of the study method. Colorectal cancers have not been associated with environmental exposures.

While the study's power to detect differences between the cohorts in rarer cancer diagnoses was limited, there is no suggestion that groups of cancers that may be associated with environmental exposures are more common in one cohort or the other.

An historical cohort design was the most feasible way to address major concerns about cancer risk in a community. The method has limitations in that lifestyle and other important factors which may be related to cancer risk could not be determined for the historical cohorts included in the study. Precise knowledge of community or individual exposures to industrial emissions remains unknown. The residential time frame chosen to indicate probable exposure in this study, 1952 - 1956, represents only minimal estimates of individuals' exposures. Indeed, individuals may have been resident in their respective communities far longer than suggested by this study's measure. The fact that these minimal estimates of peoples' exposures and the consistent findings among the study's several types of analyses suggest no increased cancer risk, supports a conclusion that there is no evidence of excess risk of cancer in one cohort or the other.

Environmental exposures to industrial emissions are less now than they were during the 1950's when this study's cohort resided in the Junction Triangle. Lack of evidence of an increased cancer risk among people living in the community at a time of relatively uncontrolled industrial emissions suggests that current risk to the community would be undetectable or non-existent.

RECOMMENDATIONS:

It is recommended:

1. That the Board of Health receive this report of the Junction Triangle Historical Cohort Cancer Study for information.

2. That no further studies to determine the risk of cancer in the Junction Triangle be planned at this time.

APPENDICES

Appendix I: Tables 5 - 13

Appendix II: Criteria for Certainty of Match Between Tax Assesment Identification and Ontario Cancer Registry

Appendix III: Glossary of Terms

TABLE 5

DISTRIBUTION OF OBSERVED CANCERS
AMONG DEFINITE, PROBABLE AND POSSIBLE MATCHES
BY COHORT AND CANCER SITE

CANCER SITE	JUNCTION TRIANGLE N	(%)	COMPARISON AREA N	(%)
Lip, oral cavity & pharynx	14	(3.8)	20	(5.8)
DIGESTIVE ORGANS:				
Oesophagus	5	(1.4)	9	(2.6)
Stomach	18	(4.9)	7	(2.0)
Colon	37	(10.1)	26	(7.6)
Rectum	12	(3.3)	14	(4.1)
Liver	2	(0.5)	3	(0.9)
Pancreas	7	(1.9)	10	(2.9)
Other digestive organs	9	(2.4)	7	(2.0)
RESPIRATORY ORGANS:				
Larynx	6	(1.6)	8	(2.3)
Trachea, bronchus & lung	65	(17.8)	55	(16.1)
Other respiratory organs	3	(0.8)	1	(0.3)
BONE, CONNECTIVE TISSUE, SKIN & BREAST:				
Bone & Connective tissues	3	(0.8)	3	(0.9)
Malignant melanoma of skin	5	(1.4)	2	(0.6)
Breast, female	42	(11.5)	43	(12.6)
GENITOURINARY ORGANS:				
Cervix	14	(3.8)	11	(3.2)
Uterus	17	(4.6)	11	(3.2)
Ovary and adnexa	10	(2.7)	6	(1.8)
Prostate	25	(6.8)	24	(7.0)
Testis, penis & other male genital organs	1	(0.3)	4	(1.2)
Bladder	16	(4.4)	19	(5.6)
Kidney and other urinary organs	13	(3.6)	5	(1.5)
BRAIN, NERVOUS SYSTEM, & OTHER ORGANS:				
Brain & central nervous system	13	(3.6)	6	(1.8)
Thyroid	4	(1.1)	0	(0.0)
LYMPHATIC AND HAEMATOPOIETIC TISSUES:				
Hodgkin's Disease	2	(0.5)	3	(0.9)
Non-Hodgkin's lymphoma	9	(2.4)	11	(3.2)
Multiple myeloma	1	(0.3)	7	(2.0)
Leukemias	4	(1.1)	10	(2.9)
OTHER SITES	9	(2.4)	17	(5.0)
TOTAL CANCERS	366	(99.8)*	342	(100.0)
NO CANCER MATCH TO REGISTRY	2,689		2,569	

* Total does not equal 100 due to rounding

TABLE 6

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS,
OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND
95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND CANCER SITE:
INCLUDING DEFINITE, PROBABLE, AND POSSIBLE MATCHES,
WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2*	RR	95% CI
MALES									
Lung	57	49.63	1.15	39	45.70	0.85	N.S.	1.34	(0.89, 2.01)
Colorectal	23	31.52	0.73	14	29.01	0.48	N.S.	1.51	(0.78, 2.93)
All Sites	201	209.25	0.96	170	192.70	0.88	N.S.	1.08	(0.88, 1.32)
FEMALES									
Lung	8	10.54	0.76	16	10.50	1.52	N.S.	0.50	(0.21, 1.17)
Colorectal	26	25.46	1.02	26	25.05	1.04	N.S.	1.00	(0.58, 1.72)
Breast	42	48.70	0.86	43	48.36	0.89	N.S.	0.97	(0.64, 1.49)
All Sites	165	168.95	0.98	172	167.35	1.03	N.S.	0.96	(0.77, 1.18)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 7

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS, OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND 95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND CANCER SITE: INCLUDING DEFINITE, PROBABLE, AND POSSIBLE MATCHES, WITH FOLLOW-UP ADJUSTED FOR LIFE EXPECTANCY

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2*	RR	95% CI
MALES									
Lung	57	44.06	1.29	39	40.25	0.97	N.S.	1.34	(0.89, 2.01)
Colorectal	23	27.28	0.84	14	24.90	0.56	N.S.	1.50	(0.77, 2.92)
All Sites	201	181.50	1.11	170	165.73	1.03	N.S.	1.08	(0.88, 1.32)
FEMALES									
Lung	8	10.20	0.78	16	10.17	1.57	N.S.	0.50	(0.21, 1.17)
Colorectal	26	24.22	1.07	26	23.76	1.09	N.S.	1.00	(0.58, 1.72)
Breast	42	47.53	0.88	43	47.17	0.91	N.S.	0.97	(0.64, 1.49)
All Sites	165	163.32	1.01	172	161.55	1.06	N.S.	0.96	(0.78, 1.19)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 8 .

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS, OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND 95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND CANCER SITE: INCLUDING DEFINITE AND PROBABLE MATCHES (POSSIBLES EXCLUDED) WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2 *	RR	95% CI
MALES									
Lung	41	50.79	0.81	25	47.45	0.53	N.S.	1.51	(0.92, 2.47)
Colorectal	18	32.28	0.56	5	30.13	0.17	6.93*	3.30	(1.23, 8.90)
All Sites	151	214.24	0.70	106	200.12	0.53	5.15*	1.31	(1.02, 1.68)
FEMALES									
Lung	8	10.84	0.74	8	10.95	0.73	N.S.	1.00	(0.37, 2.66)
Colorectal	19	26.24	0.72	13	26.15	0.50	N.S	1.46	(0.72, 2.96)
Breast	26	49.78	0.52	19	50.07	0.38	N.S	1.36	(0.76, 2.47)
All Sites	111	173.31	0.64	91	173.81	0.52	N.S	1.22	(0.92, 1.61)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 9

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS,
OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND
95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND CANCER SITE:
INCLUDING DEFINITE AND PROBABLE MATCHES (POSSIBLES EXCLUDED),
WITH FOLLOW-UP ADJUSTED FOR LIFE EXPECTANCY

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2*	RR	95% CI
MALES									
Lung	41	44.71	0.92	25	41.53	0.60	N.S	1.50	(0.91, 2.47)
Colorectal	18	27.64	0.65	5	25.67	0.19	6.87*	3.29	(1.22, 8.87)
All Sites	151	183.93	0.82	106	170.84	0.62	4.95*	1.30	(1.02, 1.67)
FEMALES									
Lung	8	10.48	0.76	8	10.56	0.76	N.S	1.00	(0.38, 2.66)
Colorectal	19	24.89	0.76	13	24.66	0.53	N.S.	1.46	(0.72, 2.96)
Breast	26	48.51	0.54	19	48.69	0.39	N.S.	1.36	(0.77, 2.46)
All Sites	111	167.19	0.66	91	167.14	0.54	N.S	1.22	(0.92, 1.60)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 10

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS,
OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND
95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND LENGTH OF EXPOSURE:
INCLUDING DEFINITE, PROBABLE AND POSSIBLE MATCHES,
WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X**	RR	95% CI
MALES									
1 Year	85	95.30	0.89	86	107.87	0.80	N.S.	1.05	(0.77, 1.41)
3 Years	48	41.17	1.17	38	44.16	0.86	N.S.	1.17	(0.76, 1.79)
5 Years	68	72.78	0.93	46	40.67	1.13	N.S.	0.88	(0.60, 1.27)
All Years	201	209.25	0.96	170	192.70	0.88	N.S.	1.08	(0.88, 1.32)
FEMALES									
1 Year	81	71.15	1.14	98	91.07	1.08	N.S.	1.00	(0.74, 1.34)
3 Years	27	32.74	0.82	40	37.32	1.07	N.S.	0.71	(0.44, 1.16)
5 Years	57	65.06	0.88	34	38.95	0.87	N.S.	1.03	(0.68, 1.58)
All Years	165	168.95	0.98	172	167.35	1.03	N.S.	0.96	(0.77, 1.18)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 11

OBSERVED (O) AND EXPECTED (E) NUMBERS OF LUNG CANCERS,
OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND
95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND LENGTH OF EXPOSURE:
INCLUDING DEFINITE, PROBABLE AND POSSIBLE MATCHES,
WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2 *	RR	95% CI
MALES									
1 Year	23	22.58	1.02	17	25.62	0.66	N.S.	1.43	(0.76, 2.68)
3 Years	15	9.88	1.52	8	10.43	0.77	N.S	1.73	(0.73, 4.08)
5 Years	19	17.17	1.11	14	9.65	1.45	N.S	0.80	(0.40, 1.60)
All Years	57	49.63	1.15	39	45.70	0.85	N.S	1.34	(0.89, 2.01)
FEMALES									
1 Year	4	4.47	0.89	11	5.75	1.91	N.S	0.44	(0.14, 1.38)
3 Years	3	2.08	1.44	3	2.36	1.27	N.S	1.05	(0.21, 5.20)
5 Years	1	3.99	0.25	2	2.39	0.84	N.S	0.31	(0.03, 3.42)
All Years	8	10.54	0.76	16	10.50	1.52	N.S	0.50	(0.21, 1.17)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 12

OBSERVED (O) AND EXPECTED (E) NUMBERS OF COLORECTAL CANCERS,
OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND
95% CONFIDENCE INTERVALS (CI), BY COHORT, SEX, AND LENGTH OF EXPOSURE:
INCLUDING DEFINITE, PROBABLE AND POSSIBLE MATCHES,
WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X² *	RR	95% CI
MALES									
1 Year	10	14.26	0.70	7	16.16	0.43	N.S	1.51	(0.57, 3.97)
3 Years	4	6.20	0.64	1	6.68	0.15	N.S	3.70	(0.41,33.10)
5 Years	9	11.06	0.81	6	6.17	0.97	N.S.	0.89	(0.32, 2.50)
All Years	23	31.52	0.73	14	29.01	0.48	N.S.	1.51	(0.78, 2.93)
FEMALES									
1 Year	12	9.68	1.24	13	12.86	1.01	N.S.	1.12	(0.51, 2.45)
3 Years	5	4.77	1.05	6	5.70	1.05	N.S	0.88	(0.27, 2.88)
5 Years	9	11.01	0.82	7	6.49	1.08	N.S	0.79	(0.29, 2.12)
All Years	26	25.46	1.02	26	25.05	1.04	N.S	1.00	(0.58, 1.72)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

TABLE 13

OBSERVED (O) AND EXPECTED (E) NUMBERS OF NEW CANCERS, OBSERVED-TO-EXPECTED RATIOS (O/E), RELATIVE RISKS (RR) AND 95% CONFIDENCE INTERVALS (CI), AMONG AGE-RESTRICTED COHORTS (BIRTHYEARS 1910-1934) BY COHORT, SEX, AND CANCER SITE: INCLUDING DEFINITE, PROBABLE AND POSSIBLE MATCHES, WITH FOLLOW-UP TO DECEMBER 31, 1982

	JUNCTION TRIANGLE			COMPARISON AREA					
	O	E	O/E	O	E	O/E	X^2*	RR	95% CI
MALES									
Lung	38	25.39	1.50	22	22.67	0.97	N.S	1.56	(0.92, 2.64)
Colorectal	14	15.37	0.91	8	13.76	0.58	N.S.	1.58	(0.66, 3.77)
All Sites	123	103.36	1.19	112	92.63	1.21	N.S.	0.99	(0.77, 1.28)
FEMALES									
Lung	6	6.81	0.88	13	6.99	1.86	N.S.	0.47	(0.18, 1.24)
Colorectal	12	12.52	0.96	15	12.90	1.16	N.S.	0.81	(0.38, 1.73)
Breast	36	32.56	1.11	36	33.13	1.09	N.S	1.01	(0.64, 1.60)
All Sites	112	101.73	1.10	118	104.02	1.13	N.S.	0.96	(0.74, 1.24)

* denotes statistical significance of the comparison between the two cohorts at the .05 level

N.S. denotes comparison between the two cohorts is Not Significant at the .05 level

APPENDIX II

CRITERIA FOR CERTAINTY OF MATCH BETWEEN TAX ASSESSMENT IDENTIFICATION AND ONTARIO CANCER REGISTRY

DEFINITE MATCH

Uncommon surname and one or more of the following:

- Match on birthdate within two years
- Treated for cancer within Metropolitan Toronto region
- Died within Metropolitan Toronto region

PROBABLE MATCH

Uncommon surname with only one source record and one or more of the following:

- Registry has only year of birth
- Registry has only partial given names

Uncommon surname with one or more of the following:

- No birthdate recorded but Registry has a birthdate which could be verified
- Birthdate match is out by more than two years
- Died outside Metropolitan Toronto region, but can use a middle name or complete birthdate to verify

Common surname and additional confirmatory identifying information, such as:

- Maiden name
- Other given names
- Complete birthdate

POSSIBLE MATCH

Common surname and one or more of the following

- Died outside Ontario
- Living outside Ontario
- Conflict in middle name or initial
- Registry has only year of birth
- A pathology report from a hospital outside Metropolitan Toronto is the only source of matching information
- Birthdate match is out by more than two years

Two or more individuals with the same name and birthdates of less than two years from birthdate identified on tax assessment file

APPENDIX III

GLOSSARY OF TERMS

CANCER REGISTRATION

A system for gathering and analyzing information on all cases of cancer and their characteristics in a given geographic area. Registries usually cover a defined population and time period.

The Ontario Cancer Registry is population-based, covering the entire resident population of the Province. Information on cases of cancer is obtained from death certificates, hospital and pathology reports, and cancer treatment centres records. The Registry includes information on all newly diagnosed cases of cancer (incidence) as well as on deaths from cancer (mortality). It has several potential uses:

- to monitor cancer trends in the population
- to assist in surveillance, early detection and control of cancers
- to help identify potential cancer-causing events or agents
- to increase basic knowledge of who is affected by cancer and, thereby, identify risk factors for the disease.

COHORT

A cohort is a group composed of members who share some common experience within a defined period of time. A cohort study identifies a group of people defined by their shared experience and follows them over time to determine the frequency of some subsequent event, such as the development of disease or death. The events to be studied may be in the future in real time, say detectible between now and the year 2000 (a prospective cohort study). Alternatively, we may wish to detect events that occurred sometime in the past, up to the current time, and perhaps continuing into the future (an historical cohort study). The latter design was chosen for this study.

Examples of cohorts are birth cohort (all people born in a given time period), marriage cohort (all people married in a given time period), or illness cohort (all people diagnosed with a specified illness in a given time period). A group of people also may be defined as a cohort because they are known or assumed to have experienced a particular exposure.

CONFIDENCE INTERVAL

A range of values within which the true value of the phenomenon being measured is thought to lie. The range between the lower and upper confidence boundaries is determined by the degree of variability in the data. A level of certainty is specified for confidence intervals. In this study, the 95% confidence interval calculated around each relative risk estimate provides upper and lower boundaries within which the true level of risk is thought to lie with 95% certainty; the probability that the relative risk lies outside these boundaries is only 5%. Confidence intervals which encompass 1.00 suggest with the specified level of certainty that the risk is not elevated.

EXPECTED FREQUENCY

The number of events, cases of death or disease that would be expected to occur in a group under study if the age and sex-specific rates of some standard population prevailed. The expected frequency is an estimate. In this particular study, age and sex-specific cancer incidence rates for the province of Ontario were applied to each of the cohorts to determine the number of cancers that would have occurred within each cohort if their cancer experience was identical to that of the province as a whole.

LATENCY PERIOD (or Induction Time)

The time interval between exposure to a disease-causing event or agent and appearance of recognizable symptoms of disease. This can be thought of as the delay between initiation of the disease-causing process and disease detection.

The latency period for most cancers is 15 to 25 years. This lengthy period before detection of disease often makes it difficult to identify a relationship between any particular exposure/agent and a disease.

OBSERVED-TO-EXPECTED (O/E) RATIO

The actual number of events or cases occurring in a group under study, divided by the number expected if some standard population rate prevailed. In this study, the actual number of cancers "observed" in each cohort (as determined by linkage to the Cancer Registry) are compared to the number that would be "expected" if the cohort had cancer rates identical to those of the provincial population as a whole. This measure serves as a means of comparing each cohort to the Province in regard to cancer experience.

PERSON-YEARS

A measure which combines people and time. The length of time each person contributes to a calculation of risk is determined by the number of years he/she is actually observed (or lives) during the period covered by the study. (e.g., One person followed for 10 years contributes 10 person-years; 10 people followed 1 year each also contribute 10 person-years.)

In this historical cohort study, each person's contribution is calculated as the number of years between 1964 and the year of their diagnosis if cancer was detected, or between 1964 and the end of 1982 if he/she did not develop cancer.

RELATIVE RISK

A measure of the risk of disease or death among a group exposed to some factor presumed to contribute to disease or death, compared to an unexposed group. Relative risk is calculated as the ratio of the incidence rate of disease or death among those exposed, to the incidence rate among those unexposed. A relative risk about 1.00 indicates no excess risk, while below 1.00 suggests reduced risk and above 1.00 an elevated risk compared to the unexposed group. In this study, relative risk provides an estimate of any excess risk of cancer experienced by residents of Junction Triangle.

REFERENCES

1. Metropolitan Toronto Planning Department. Industrial Land Use Data, 1976.

2. Ministry of the Environment, "Special Report on the Junction Triangle", not dated.

3. Medical Officer of Health, Chronology of Involvement in the Junction Triangle, November 28, 1984, Board of Health, City of Toronto.

4. Junction Triangle Anti-Pollution Committee, Submission to the Toronto Board of Health, Toronto, 1979.

5. Dewar E, "Inside the Junction Triangle", Toronto Life, July:33-63, 1984.

6. Spitzer WO, A Study of the Health Status of Residents of the Junction Triangle Toronto, City of Toronto, Department of Public Health, September 1984.

7. Medical Officer of Health, Junction Triangle Health Assessment Program, November 19, 1985.

8. Medical Officer of Health, Junction Triangle Historical Cohort Cancer Study, November 27, 1984, Board of Health, City of Toronto.

9. Murdie RA, Junction Triangle Historical Cohort Cancer Study: Recommendations to the City of Toronto Department of Public Health on Census Tracts considered to be (a) Most Demographically Similar to the Junction Triangle and (b) the Larger Junction Area but Minimally Exposed to Industrial Emissions. Department of Geography, York University, January 1985.

10. Metropolitan Toronto Planning Department. Industrial Land Use Data, 1958.

11. Walter SD, Determination of significant relative risks and optimal sampling procedures in prospective and retrospective comparative studies of various sizes. American Journal of Epidemiology 105:387-397, 1977.

12. Coleman M, Douglas A, Hermon C, Peto J, Cohort study analysis with a FORTRAN computer program. International Journal of Epidemiology 15:134-137, 1986.

13. Berry G, The analysis of mortality by the subject-years method. Biometrics 39:173-184, 1983.

14. Kleinbaum D, Kupper L, Morgenstern H, Epidemiologic Research. California: Lifetime Learning Publications, 296-299, 1982.

15. Dominion Bureau of Statistics Cat. No. 84-510. Canadian Life Tables, 1955-1957. Ministry of Industry, Trade, and Commerce. Ottawa: 1960.

16. Toronto Department of Public Health, Community Health Survey 1983, Unpublished data.

17. Bailer JC, Ederer F, Significance factors for the ratio of a Poisson variable to its expectation. Biometrics 20:639-643, 1964.

18. Schottenfeld D, Fraumeni JF Jr, Cancer Epidemiology and Prevention. Toronto: W.B.Saunders Company, 1982.

19. Ontario Cancer Registry, 1986, Personal communication.

20. Doll R, Peto R, The Causes of Cancer. New York: Oxford University Press, 1981.